INTERNATIONAL JOURNAL OF BUSINESS, 30(4), 2025 ISSN: 1083-4346

Informing Private Equity Selection for Limited Partners using Machine Learning

Daniel R. Cavagnaro [a], Yinfei Kong [b], Yingdi Wang [c, *]
[a,b] *Department of Information Systems and Decision Sciences, California State University, Fullerton*
[c] *Department of Finance, California State University, Fullerton*
[a] *dcavagnaro@fullerton.edu*
[b] *yikong@fullerton.edu*
[c] *yingdiwang@fullerton.edu.*

ABSTRACT

We evaluate whether limited partners can use machine learning on structured data to help identify top-performing private equity funds. We train and test six supervised learning algorithms (linear probability, regularized and additive, and tree-based models) on a sample of 1,402 venture and buyout funds raised between 1990 and 2011. We find that using these models to select investments can generate economically large returns. While the models vary in their ability to identify top performers accurately, they generally outperform benchmark heuristics in out-of-sample tests. The strongest gains are found for buyout funds and for the combined sample. Results for venture funds show less pronounced benefits. We further find that machine learning considers multi-dimensional interaction terms, rather than single fund or firm characteristics, important for predicting returns. Overall, our results suggest that machine learning can serve as a tool to aid limited partners in their decision-making process, particularly in buyout fund selection.

JEL Classifications: G11, G23, G24

Keywords: private equity, machine learning, institutional investors

https://doi.org/10.55802/IJB.030(4).002

I. INTRODUCTION

Due to high early returns, private equity, as an asset class, has experienced dramatic growth. Between 2013 and 2015, private equity raised over $300 billion. In 2021, global fundraising reached $1.2 trillion[1]. Many institutional investors have allocated a significant portion of their assets to private equity, making returns from this asset class economically meaningful to those institutions and their beneficiaries. As a result, a longstanding problem in private equity has been to determine if and how these limited partners (LPs) can predict the performance of new funds.

One approach to predicting fund performance is to use past returns on funds raised by the same general partnership (GP). However, past returns are not always a reliable indicator of future performance. They are noisy for identifying skilled GPs (Korteweg and Sorensen, 2017), and persistence of returns has declined substantially post-2000 for buyout funds (Harris et al., 2020; Braun et al., 2017). In addition, GPs market their next fund based on interim valuations of their current fund, and firms time their fundraising to match interim valuation peaks (Barber and Yasuda, 2017). These interim performance measures have little power to predict ultimate returns (Jenkinson, Sousa, and Stucke, 2013).

Related work has utilized regression analyses to identify characteristics that correlate with fund returns (e.g., Phalippou and Zollo, 2005; Diller and Kaserer, 2009). However, identifying correlates is a different problem from predicting the ultimate performance of funds. The prediction problem is one for which regression analyses may not be well-suited due to the structural limitations and auxiliary assumptions imposed by regression. The prediction task is particularly challenging in part because there is so much variation in the quality and returns of funds. Between 1990 and 2011, the difference in average internal rate of return (IRR) between top- and bottom-quartile funds is close to 20 percentage points[2]. In addition, the complex relationship between fund/firm characteristics and predicted performance may not be as simple as a tradeoff summarized by coefficients estimated using regression models. Furthermore, regression coefficients often oversimplify the complex relationships between fund characteristics and performance. Imposing such a structure on the data when it is not warranted can lead to biased estimates and inaccurate predictions.

As regressions are not well-equipped to make ex-ante predictions in private equity, it remains unclear how LPs can use widely observable quantitative data to identify top-performing funds. In this paper, we move beyond the traditional regression methods and consider whether machine learning can aid LPs' decision-making by predicting top-performing funds using observable structured data. Unlike regression models, machine learning models are specifically intended to maximize predictive power, and they can be used with less concern about the validity of underlying structural assumptions. This approach also allows for more complex, nonlinear, interactive relationships among fund, GP, and market characteristics that may be important for ex-ante predictions.

Recent performance predictions in financial economics have successfully utilized machine learning and other technical approaches (e.g., Gu et al., 2020; Gupta and Nieuwerburgh, 2021; Kaniel et al., 2022; Brown et al., 2023; Fragkiskos et al., 2025).

---

[1] Bain & Company Global Private Equity Report "The Private Equity Market in 2021: The Allure of Growth."

[2] Results are obtained from taking the average IRR of funds covered by Preqin.

https://doi.org/10.55802/IJB.030(4).002

Unlike those studies, our aim is not to measure abnormal returns or replicate portfolios. Instead, our goal is prediction and classification of top-performing funds based on characteristics that all LPs can observe. Doing so broadens LPs' portfolios, reduces search costs, and improves returns, all of which significantly impact the beneficiaries of these institutions.

We conduct our analyses on a large sample of venture and buyout funds raised between 1990 and 2011 using variables that capture characteristics at the fund, GP, and general private equity levels. Due to the noise in private equity returns, we do not predict funds' percentage returns. Instead, we use those observed characteristics to identify top-performing funds.

We consider several machine learning algorithms to classify fund performance, including two logistic models[3] and four supervised learning models. We train each algorithm to identify which funds will be in the top quartile of IRRs, relative to funds of the same type and vintage year[4]. Hereafter, we refer to top-quartile funds as C_1, and all other funds as C_0. We use cross-validation to evaluate how well each algorithm can correctly identify such top-performers. Specifically, in each of 100 iterations, we randomly select 85% of our full sample on which to train each algorithm with true classifications. We then use the trained algorithms to predict the classifications of the remaining 15%. Besides this full-sample split, we also consider two alternative splits. The first is a partial-sample split using only the data after the year 2000 to account for the changing private equity landscape after 2000 (Sensoy et al., 2014). The second is a chronological rolling-window split to account for potential look-ahead biases and assess the stability of results over time. Results are discussed later.

We use several metrics to evaluate these classifications generated by the algorithms. The first is based on accuracy, sensitivity, and specificity. Accuracy measures the proportion of funds correctly classified. Sensitivity and specificity separately measure the proportion of true C_1 and C_0 funds correctly classified. The second metric is based on how well each algorithm can discriminate between true C_1 and true C_0 funds. Lastly, to give an economic interpretation to the numbers, we consider the average IRRs of the funds predicted to be C_1 by each algorithm. To adjust for time trends of IRRs, we also calculate demeaned IRRs by subtracting the average IRRs of funds within the same type and vintage year.

We also compare the results of our algorithms with those based on two other strategies for selecting funds: *return chasing* and *random selection*. Return chasing selects follow-on funds of top-quartile funds for each fund type and vintage year. It captures the extent of return persistence. While random selection does not represent the strategy of a typical LP, it provides a straightforward benchmark.

We find that, compared to both random investing and return chasing, machine learning algorithms are better at identifying true C_1 funds when considering buyout and venture funds together. However, there is variability among algorithms in their ability to classify funds. We also find that funds predicted as C_1 by any algorithm tend to generate economically large returns. The average IRR of all funds predicted to be C_1 is higher than 30 percentage points for three algorithms. In addition, the hypothetical returns from most

---

[3] We do not use logit models to estimate coefficients. Rather, we use them to classify funds in the same way as we use GAM, GLMnet, random forest, and XGBoost. Therefore, logistic models, though often used in regressions, are used as a machine-learning algorithm as well.

[4] We also repeat the same process for classifying bottom-quartile funds. Results reported in Appendix II.

https://doi.org/10.55802/IJB.030(4).002

of the tested algorithms are also larger than those from return chasing and are much higher than the returns from random investing. The general pattern of machine learning doing well with variability in performance holds when considering buyout funds separately. However, the advantage of using machine learning for venture funds is less obvious. The same results hold when we only use the post-2000 sample period. These results are consistent with the existing findings that venture funds are noisier (Korteweg and Sorensen, 2017) and continue to show persistence (Harris et al., 2020).

Our results are robust to testing in different market conditions, over rolling windows, removing any look-ahead bias, and considering the risk of the funds. We also use the same algorithms to identify bottom-quartile funds for LPs to avoid. Besides the pre- and post-2000 split, we use a rolling window by training the algorithms on earlier samples and testing in the following vintage year. We do this for each vintage year starting in 2007 to ensure that the predictions remain stable over time. We further remove variables (e.g., lag IRR and fund size) that may not be available at the time of fundraising. The results are not meaningfully different.

To assess whether using machine learning would produce riskier fund selections, we examine the average C_1 fund IRRs across 100 cycles. We find that the variances of all funds and venture funds produced by machine learning algorithms are much lower than those from return chasing, while buyout funds show only slightly higher variances compared to return chasing. The comparisons do not suggest riskier funds from machine learning. We also find that the algorithms can help identify low-performing funds to avoid. This can further reduce the concern over investing in risky funds.

One natural question that arises from predictions is whether there are characteristics that are associated with better performance. To that end, machine learning algorithms do not produce significance of estimated coefficients as a regression model does. However, we can identify a set of characteristics that are the most important in training the data for some algorithms, which in turn tells us which characteristics are considered important. A striking result from the list of top characteristics is that almost all the top variables picked by the algorithms are interaction terms. This is in direct contrast to the existing literature, which often focuses on single characteristics such as fund size and preceding fund returns. This result highlights the complexity of evaluating private equity returns. It suggests LPs should look beyond past returns. Instead, they should evaluate how combinations of interrelated factors like GP experience, historical performance, fund size, and market conditions interact.

Our finding that machine learning can help classify top-performing funds is related to the work from Braun et al. (2023), which uses natural language processing techniques and machine learning to predict fund performance. However, our paper differs from Braun et al. (2023) in several regards. First, this paper uses structured data that all LPs would be able to access, while Braun et al. (2023) focus on proprietary qualitative textual data from private placement memoranda (PPMs). They find that machine learning is remarkably effective in using these qualitative data. Second, we focus on a different set of performance criteria. In this paper, we aim to identify top-quartile funds and evaluate the classifications of those top performers. In contrast, Braun et al. (2023) define success based on the median total value to paid-in of funds.

This paper makes several contributions to the literature. First, machine learning has been successfully used for prediction and classification in many fields as well as other areas of finance (e.g., Gu et al., 2020; Erel et al., 2021; Li et al., 2021; Kaniel et al., 2022).

In private equity, machine learning can be used by venture capital GPs to predict post-money valuation and success (Ang et al., 2021) and broaden the scope of venture capitalists' investments (Davenport, 2022; Lyonnet and Stern, 2025). Here, we take on the perspective of LPs and show that these algorithms can be useful in fund selection. However, unlike studies in other fields that document a clear advantage of machine learning, we also find that the benefits of using this approach are less pronounced for private equity funds, where information is limited and complex. This result is complementary to those documented in Braun et al. (2023), which highlights the value of qualitative information in predicting fund performance.

Second, the literature has shown that selecting better investments is economically meaningful for LPs, and that institutions vary in their private equity returns (Lerner et al., 2007; Hochberg and Rauh, 2013; Sensoy et al., 2014; Cavagnaro et al., 2019). We demonstrate that using machine learning can help some LPs utilize hard information in their decision-making.

While our results highlight the potential of using machine learning in private equity, they also show that algorithms cannot replace LPs' decision-making. The predictive power and economic benefits gained by deep learning are not equal across all fund types. Hochberg et al (2014) point out that LPs learn valuable soft information about GP skill from investing in the GPs' past funds. Braun et al. (2023) further demonstrate the significance of analyzing qualitative information. Therefore, this machine learning method should be used to supplement information that LPs already have in estimating the ultimate performance of a new fund.

The rest of the paper is organized as follows. Section II reviews related literature and offers economic intuition on the characteristics that can be related to fund performance. Section III describes the sample and summary statistics. In Section IV, we formulate the empirical approach and lay out the results. In Section V, we describe the characteristics that are related to fund performance, and Section VI concludes.

II. LITERATURE REVIEW

Previous studies on private equity performance imply that returns can be driven by factors related to the specific funds, GPs, and the private equity industry in general. However, there is no clear consensus on how to use structured data to predict private equity performance.

Much of the literature's effort on predicting returns focuses on preceding fund performance. Kaplan and Schoar (2005) first documented that returns persist strongly across subsequent funds of a partnership. However, later studies document a decline in persistence. Chung (2012) shows that the decline occurs after the first follow-on fund. Harris et al. (2020) find little evidence of persistence among buyout funds raised after the year 2000, while venture funds' returns continue to show persistence. Braun, Jenkinson, and Stoff (2017) also show that buyout persistence has substantially declined.

In addition to past performance, the literature finds that returns can be related to fund size and GPs. Kaplan and Schoar (2005) argue that returns and fund size follow a concave relationship and that they differ with GP experience. Ewens and Rhodes-Kropf (2015) show evidence of persistent performance at the partner level. Korteweg and Sorensen (2017) point out that the structure of private equity returns does not make an ideal setting for running regressions. Using Bayesian estimation, the paper finds long-

https://doi.org/10.55802/IJB.030(4).002

term persistence consistent with skill, but past returns are noisy. Furthermore, Jones and Rhodes-Kropf (2003) show that returns are driven by idiosyncratic risk, although the literature does not offer a clear measure of risk.

Returns can be related to the general private equity market as well. Gompers and Lerner (2000) document a money-chasing-deals phenomenon that negatively relates returns to fund inflows. Diller and Kaserer (2009) test for drivers of fund performance and confirm the importance of fund flows in explaining returns, in addition to measures that represent GP skill and risk. Phalippou and Zollo (2005) also studied the drivers of performance using different variables and found that performance is procyclical.

These aforementioned papers highlight the difficulty of evaluating fund performance: returns are affected by many different variables, some of which are difficult to measure and can interact with each other, and information is noisy for reliably identifying top performers. Studies that do not use a regression approach have had some success in evaluating fund performance (e.g., Brown et al., 2023). More recently, Braun et al. (2023) argue that the use of using machine learning algorithms and natural language processing techniques using qualitative data can help predict fund performance. They construct textual data using PPMs from a large institutional investor in Europe. They find that the three algorithms they select are effective in using qualitative information to predict fund performance, while traditional measures such as fund size and past performance are not good predictors of performance.

III. SAMPLE DESCRIPTION AND SUMMARY STATISTICS

Our sample consists of buyout and venture fund information collected from Preqin, a leading source of data for private equity returns. Preqin's data is collected through regulatory filings, press releases, news, and websites, as well as Freedom of Information Act requests to LPs and voluntary reporting by GPs and LPs. The database covers information on more than 19,000 private funds.

From all buyout and venture funds in Preqin, we delete funds with missing vintage year, fund size, and fund IRR, which we use to measure returns[5]. We also delete funds smaller than $5 million and restrict our sample to the period from 1990 to 2011, to give it sufficient time to observe fund returns[6]. This process gives us a total of 2,377 funds. Finally, since our tests rely on preceding fund IRR, we delete all observations with no preceding fund returns. This includes all first-time funds and some higher sequence funds. Our final sample includes a total of 1,402 funds, out of which 776 are buyout funds and 626 are venture funds.

Table 1 shows the mean, median, first, and third quartile values of the observed characteristics in our sample[7]. We also divide the full sample of funds into buyout and venture funds separately. The statistics on IRR are in line with those shown in existing literature. On average, funds in our sample earn 13.94 percent IRR. There is much variation in IRR, though, as the Q1 value for IRR is 3.2 percent while Q3 value is close

---

[5] IRR is a typical net-of-fees performance measure for private equity funds and is directly reported by Preqin. As the data we have do not include any information on cash flows, we cannot compute public market equivalent (PME) or net asset values for returns.

[6] Private equity funds take 10 years to realize returns on average, with some funds taking 10-12 years.

[7] GP location, fund region focus, and demeaned IRRs are also included in our tests but not reported in this table because they are either character variables or the summary statistics would not be meaningful.

https://doi.org/10.55802/IJB.030(4).002

to 20 percent. IRRs also vary by fund type. The average, median, Q1, and Q3 values of IRR of venture funds are all lower than those of buyout funds.

We also include additional fund-level variables, including lag IRR, fund size, industry indicators,[8] as well as GP-level variables that capture experience and long-term persistence, and finally, variables at the private equity industry level related to total inflows and market competitiveness. A detailed description of all the variables used is included in Appendix I. Overall, buyout funds are much larger than venture funds, and they show higher lag returns and GPs' average returns. However, there are more venture capital funds in the IT or healthcare industry.

IV. EMPIRICAL APPROACH AND RESULTS

A. Methodology

We consider six machine learning algorithms for binary classification. The first two are logistic regression models. The first of these, which we call *Logit 0*, uses demeaned preceding fund IRR ("lag IRR") as the only predictor. We include this model as a benchmark to demonstrate how far one can get using lag performance alone, since much of the focus of past studies is on lag performance. The other logistic regression model, which we call *Logit*, uses the full list of independent variables (see Appendix I). The other four algorithms are common, supervised learning models, which we implement with publicly available R packages. *GAM* fits a generalized additive model to the data, which extends the General Linear Model (GLM) to allow for nonlinear relationships. *Glmnet* implements Lasso regression. Finally, *random forest* and *XGBoost* are decision tree algorithms.

We use a train-test split to evaluate each algorithm. This technique entails randomly selecting a subset of the full sample as the training set. To ensure enough data in the training set, we use 85% of the full sample for training and the remaining subset as the validation set. We fit each model to the training set to obtain classification probabilities for the funds in the validation set. We then evaluate the performance of each algorithm by comparing its classifications for the validation set with the actual performance of the funds. To ensure our results are robust and not skewed by a single sample, we repeat this train-test process 100 times using different random splits.

Our main analyses rely on a random split, instead of a chronological one, because the private equity industry has changed over time (Sensoy et al., 2014). As a robustness check, we consider two different chronological splits of the sample as well.

B. Comparisons of Predictions Across Algorithms

We begin by assessing each algorithm's accuracy (i.e., proportion of all funds classified correctly), sensitivity (i.e., proportion of true C_1 funds classified correctly), and specificity (i.e., proportion of true C_0 funds identified correctly). To do this, we must first

---

[8] The fund industry in Preqin is very finely defined and has over 1,000 combinations. Therefore, we group funds into broad industry classifications based on whether they are in information technology (IT) (e.g., IT, technology, software, internet) or healthcare (e.g., healthcare, life science, pharmaceuticals). If an IT- or healthcare-related industry appears in the fund's description, then we assigned it to IT or healthcare, respectively. There are funds that are in both industries or neither one of them.

https://doi.org/10.55802/IJB.030(4).002

convert the classification probabilities produced by the algorithms into binary classifications using cutoffs. For example, with a cutoff of 0.9, we treat classification probabilities greater than 0.9 as predictions of C_1, and all classifications less than 0.9 as predictions of C_0. There is no standardized cutoff to use. Therefore, we consider several cutoffs, ranging from 0.1 to 0.9 in increments of 0.1, and report separate results for each. These represent different thresholds that can help with making decisions. A lower cutoff identifies more top-quartile funds but includes more false positives, while a higher cutoff represents a more conservative decision with fewer fund selections.

Panel A of Table 2 reports the average accuracy, sensitivity, and specificity of each algorithm, at each cutoff, across 100 train-test splits. This panel highlights the statistical measures of using algorithms to identify top-quartile funds. The first column shows the cutoff on classification probabilities under which funds are predicted C_1. To be brief, we only report cutoffs of 0.1, 0.2, 0.3, 0.4, and 0.5 in this table. Results for the rest of the cutoffs are similar, and therefore, are reported in Table AI.2 in Appendix I. Column *Predicted C_1* shows the percentage of identified true C_1 funds under the corresponding cutoff. More lenient cutoffs (e.g., 0.1) result in more (or even all) funds being predicted C_1.

Although accuracy may be the most intuitive of the measures shown in Table 2, a higher accuracy does not necessarily indicate better predictive performance. Since approximately 75% of funds in each validation set are true C_0, an algorithm that predicted all funds to be C_0 would achieve 0.75 accuracy without correctly identifying any of the C_1 funds (i.e., with a sensitivity of zero). A useful algorithm should correctly identify at least some of the true C_1 funds without also incorrectly classifying many true C_0 funds as C_1. Therefore, what we seek in Panel A of Table 2 is an algorithm that achieves high marks in each of accuracy, sensitivity, and specificity. However, there is a trade-off between sensitivity and specificity related to the cutoff: as the cutoff increases, fewer funds are predicted C_1, so sensitivity goes down, and specificity goes up. The appropriate balance between sensitivity and specificity may depend on the LPs' goals. Do they want to identify as many C_1 funds as possible, or identify just a few C_1 funds with very high confidence? For the latter, a high cutoff may be more appropriate. All algorithms except for Logit reach a specificity of at least 98% with a cutoff of at least 0.5, while still correctly identifying several true C_1 funds.

To provide a benchmark for the algorithms' results, we consider two naïve investment strategies: random investing and return-chasing. The random investing strategy selects, uniformly at random, 25% of the funds in the validation set to classify as C_1 and the remaining 75% as C_0. The return-chasing strategy considers the IRR of the preceding fund for each fund in the validation set and classifies the 25% of funds with the highest lag IRR as C_1.

We report the accuracy, sensitivity, and specificity of these benchmark strategies in Panel B of Table 2[9]. To compare these with the statistics produced by machine learning, we focus on the machine learning cutoff of 0.3 shown in Panel A of Table 2. This cutoff matches more closely with the selection rates of the benchmark strategies, thus providing a better comparison. For most algorithms, this cutoff results in a

---

[9] The random investing strategy correctly identifies 25% of true C1 funds, 75% of true C0 funds, and 62.5% of funds overall. These are expected percentages that we calculate analytically, rather than in split-test simulation.

percentage of funds predicted C_1 close to the 25% used for both benchmarks. The results show that the algorithms do better at identifying true C_1 funds at a cutoff of 0.3. Both sensitivity and predicted proportions of true C_1 funds produced by most algorithms are higher than those of the benchmarks. We find similar results when we separate all funds into buyout and venture funds. Overall, the statistics shown in this table suggest that machine learning algorithms can outperform benchmarks, particularly for buyout funds.

C. Discrimination between True C_1 and True C_0 Funds

Stepping away from dichotomizing cutoffs, Figure 1 compares each algorithm's probabilistic evaluations of true C_1 and true C_0 funds. This figure reinforces the result that machine learning algorithms are better at identifying top performers, particularly for buyout funds. The High (Low) box represents the probabilities assigned to true C_1 (C_0) funds. Intuitively, an algorithm distinguishes the funds well if it tends to assign higher probabilities to the true C_1 funds than to the true C_0 funds. Thus, we assess the performance of each algorithm by the degree to which the high box is shifted higher than the Low box.

For all funds, neither logit model shows much difference between the High and Low boxes[10]. In contrast, the more complex algorithms all have visible differences between the high and low boxes. However, random forest shows the largest difference, implying that the algorithm does best in differentiating fund classifications correctly.

When we separate funds by type, we see different patterns for buyout and venture funds. In general, the differences are larger for buyout funds than they are for venture capital funds, suggesting that the algorithms can better classify buyout funds than venture capital funds. This is in line with the notion that venture fund returns are noisier. Overall, it seems that Glmnet shows the largest difference for buyout funds, while random forest performs the best for venture funds.

For each cycle, we also conduct a two-sample t-test of assigned probabilities between true C_1 and C_0 funds in the test data. A small p-value means that the algorithm assigns higher probabilities to true C_1 funds than to true C_0 funds, on average. In Figure 2, we report the p-values of those t-tests from 100 cycles.

The results in this figure support the patterns shown in Figure 1. Glmnet, XGBoost, random forest, and GAM produce much lower p-values than those from the two logit models. We also see a difference when separating funds by type. For buyout funds, Glmnet has the lowest p-values, while random forest shows the lowest p-values for venture funds. In addition, the p-values are generally higher for venture capital funds than they are for buyout funds.

Overall, Figures 1 and 2 contain more informative assessments of the quality of machine learning classifications. We find that the algorithms are better able to classify buyout funds than venture capital funds. However, there are differences among the algorithms. The more complex machine learning algorithms can better distinguish C_1 funds from C_0 funds compared to the two Logit models. In particular, Glmnet does better for buyout funds, while random forest works best for venture capital funds.

---

[10] For the full Logit model, there are too many variables relative to the sample size, resulting in probabilities of either 0 or 1 for many funds.

D. Economic Significance of Differences in Predictions

One consideration for using machine learning to predict top-performing funds is whether using it would make any economic difference in performance. We expect the machine learning algorithms to generate larger returns for LPs if they can better identify top performers. Therefore, we evaluate the performance of each algorithm and benchmark in terms of the hypothetical average IRR of predicted C_1 and C_0 funds. We expect the machine learning algorithms to generate larger returns for LPs if they can better identify top performers.

For each cycle, we compute the average IRR of investing in predicted C_1 and C_0 funds following the algorithms. We then compute the average IRR for these investments across 100 cycles. We also produce benchmark average IRRs following the two investment strategies shown in Table 2. In the case of random investing, the average IRRs would be the same as the average from the sample. For return-chasing, we produce the average IRRs using follow-on funds of top quartile funds, given the fund type and vintage year.

Table 3 reports the results for all algorithms and benchmarks. These results can be used to evaluate the economic significance of machine learning. Our top-quartile classifications are based on demeaned IRR. However, to better interpret the numbers, Table 3 reports average IRR (Column *Average IRR*) in addition to the demeaned averages (Column *Demeaned IRR*). Because funds being classified as top quartile change with different cutoffs, we use the cutoff that maximizes the average demeaned IRR of the predicted C_1 funds. We call this *Optimal Cutoff*.

Results in Table 3 show that the algorithms can generate high returns. For all funds, most algorithms generate an average IRR of over 20 percent for C_1 funds, higher than the Q3 value reported in Table 1. Those numbers are also higher than the average IRRs generated by random investing and return chasing. Results are similar when we evaluate demeaned averages.

In addition, most algorithms show considerable differences in the average IRRs of C_1 and C_0 funds. This suggests that machine learning algorithms can more accurately classify C_1 and C_0 funds. Logit 0 does surprisingly well for all funds, generating the highest IRR, but at a cutoff where it only predicts 0.011 of funds to be C_1.

When we separate funds by type, we find that the results for buyout funds are similar to those for all funds. Except for Logit, all algorithms generate average IRRs exceeding the Q3 value for buyout funds shown in Table 1. Most algorithms also outperform random investing and return chasing. In particular, random forest and Glmnet show the highest average C_1 IRRs and the highest difference between C_1 and C_0 IRRs.

Similar to the p-values shown in Figure 2, we see less advantage of using machine learning when we evaluate venture funds. The average IRRs of predicted C_1 funds are still high, and most exceed those of random investing. However, fewer algorithms produce average IRRs that are higher than those from return chasing. We also see less difference between the average IRR of predicted C_1 and C_0 funds. The lack of an obvious advantage for venture funds could be due to their returns still persisting (Harris et al., 2020), and as a result, machine learning does not produce as much of an economic benefit when compared to return chasing.

To deal with extreme performers and outliers, we re-calculate hypothetical average IRRs using the average returns of the entire fund class. All true C_1 (C_0) funds are assigned the average IRR of all top-quartile (non-top-quartile) funds within the same vintage year and fund type. These are shown in Table 4, which reassesses economic performance less influenced by potential outliers.

For all funds and buyout funds, the main results are similar to those in Table 3: machine-learning algorithms generate higher IRRs and higher differences in IRR of C_1 and C_0 funds. However, Table 4 also shows that Glmnet generates high returns using average IRRs of the fund class. This suggests that Glmnet does well in identifying C_1 funds, even though it does not necessarily classify the highest-performing funds. For venture funds, we do not see any obvious economic significance of using machine learning over the conventional wisdom of investing.

Sensoy et al (2014) show that the private equity landscape has changed over time, and persistence in buyout funds has declined post-2000 (Harris et al., 2020; Braun et al., 2017). Therefore, we rerun our tests using the 2000 to 2011 sample period. Results are reported in Table 5. To be brief, we only report results using demeaned IRR.

Not surprisingly, compared to the full sample period, we find a decline in the average IRR of predicted C_1 funds following each of the two benchmark strategies. However, the results continue to show that most algorithms outperform the benchmarks for all funds and buyout funds. They generate higher average IRRs of C_1 funds and show a larger difference between C_1 and C_0 funds. Results for venture funds do not show much advantage in the algorithms over the benchmarks.

Overall, our results suggest that we can use machine learning algorithms to help predict funds' performance. These algorithms yield better classifications than what chasing returns and random investing produce. In addition, utilizing machine learning can produce high returns, especially for all funds and buyout funds. The economic gains of using machine learning are large, especially in the context of Cavagnaro et al. (2019), which shows that a one standard deviation increase in skill leads to a one to three percent increase in IRR. However, we also find that the advantage of machine learning over the conventional wisdom of investing is not pronounced for venture funds, as their returns continue to show persistence.

It is worthwhile noting that machine learning relies heavily on data availability, and private equity is known for limited data. Our tests only pertain to the fund-level data in Preqin. We also do not have information on cash flows that can be used to compute public market equivalents. Presumably, LPs who have more detailed information on funds and their GPs can better utilize this method to predict returns. In the following subsection, we address other concerns that may arise from the sample.

E. Robustness Checks

We end our sample period in 2011 to allow time to observe returns. This may raise concerns over whether results are stable enough to be applied to recent data. Our tests also assigned funds randomly to the training and validation sets in the sample period. This means that some training sets may include funds that were raised after other funds in the validation set. To ensure results remain unchanged over time and remove any potential look-ahead bias, we use a rolling window split of the data instead. For each vintage year, we train the algorithms on all the data for funds raised before that year, and use funds raised in that year as the validation set. We started our validation data in 2007 to have a large enough training set. We further remove all information that may not be available to the LP at the time of fundraising. These include lag IRR (both raw and demeaned), fund size, and fund growth variables.

Results for each validation year are reported in Table 6. Logit 0 is excluded from this table as a result of removing lag IRR. We only report accuracy, sensitivity, and specificity for a cutoff of 0.2 to be brief. We find that for most algorithms, the accuracy of the algorithms' classifications reduces due to a reduction in specificity. The same results hold when we separate funds into buyout and venture funds. This reduction in accuracy is not surprising, as we removed all variables related to fund size and lag IRR.

Another potential concern is that the algorithms might be picking out riskier funds. If that were true, then this method might not be well-suited for risk-averse LPs. There is no straightforward way to adjust for risk in private equity. To alleviate this concern about risk, we examine the variance of IRRs of predicted C_1 funds[11]. As a reference point, we use the variance of funds predicted by return chasing.

Results reported in Table 7 show that for all algorithms, the variances of predicted C_1 IRRs are much lower than that of return chasing for all funds and venture funds. For buyout funds, the variances of IRRs produced by the algorithms are slightly higher than those of return chasing. All methods also show higher variances for venture funds than buyout funds. This is due to higher variation among venture returns. The results for demeaned IRRs are similar as well.

Overall, the results do not support the argument that the algorithms pick riskier funds compared to the benchmark. Furthermore, LPs can address the concern about risk by utilizing different cutoffs for fund selection and using algorithms to identify funds to avoid. More risk-averse LPs can choose a higher cutoff, indicating a higher probability of the fund performing well. However, the trade-off is fewer identified C_1 funds. We also test whether the algorithms can correctly identify low-performing funds for LPs to avoid. Results reported in Appendix II confirm that the algorithms can better identify bottom-quartile funds compared to the benchmarks.

V. CHARACTERISTICS RELEVANT FOR PREDICTING FUN EMPIRICAL APPROACH AND RESULTS

Machine learning helps identify top funds, but interpreting how individual variables contribute to the result remains difficult. To get a sense of which characteristics are the most important in predicting C_1 funds, we identify the variables that the fitted algorithms "weight" the most heavily. In other words, they do not represent causal relationships but offer insight into the patterns or combinations that the algorithms rely on to classify funds. We focus on the variables identified by random forest and Glmnet, since we find these two algorithms work best. For Glmnet, we rank characteristics based on the frequency of the variable being selected across 100 cycles, and for random forest, we use the average importance score across all cycles. Characteristics with higher selection frequencies or average importance scores are ranked higher.

Table 8 lists the top 20 characteristics in each algorithm for all funds. This table provides a picture of the most influential predictors considered by algorithms. While the algorithms do not agree on all the top-20 variables, they show a clear picture of the non-linear nature of private equity performance that cannot be explained by any single historical variable. Almost all variables in this table are interaction terms that include fund size, focus, GP age, and past returns that extend beyond the preceding fund return.

https://doi.org/10.55802/IJB.030(4).002

For example, both algorithms prioritize the interaction between GP age and historical returns. This suggests that while past returns are informative, their predictive value depends on the GP's age. This result is relevant to a learning framework such as the one outlined in Chung et al. (2012), which uses performance as a signal of GPs' ability. This interaction term shows that the informativeness of returns as a signal for GPs' ability differs by GPs' age.

Table 8 also reveals that the algorithms weigh the interactions almost as much as preceding fund returns. In terms of both selection frequency and average importance score, Lag demeaned IRR does not seem drastically different from some other interaction terms that capture fund size, GP historical performance, and general private equity market conditions. The importance of these interaction terms is in line with the findings that fund size, GP experience, and fund inflows are all significantly related to returns (e.g., Gompers and Lerner, 2000; Phalippou and Zollo, 2005; Kaplan and Schoar, 2005; Diller and Kaserer, 2009; Korteweg and Sorensen, 2017). As Diller and Kaserer (2009) point out, returns are determined by multiple aspects of the fund, GP, and the private equity market. Our machine learning results echo the importance of this interplay of these characteristics: the models use a combination of fund, GP, and market characteristics to identify top performers.

This means when evaluating private equity funds, LPs cannot focus only on past performance, especially with mega funds being more prolific in recent years. Different characteristics, including size, affect returns simultaneously. Considering aspects such as GP age, whether the GP location and fund focus are in the same region, and fund growth can potentially result in higher returns.

VI. CONCLUSION

In this paper, we show that machine learning algorithms can be used to help LPs identify top-performing buyout funds and venture funds to a lesser extent. We do this by first comparing the classifications of several popular linear probability, regularized/additive, and tree-decision algorithms with those of two benchmark strategies: random investing and return-chasing. Then, we compare the economic significance of utilizing machine learning. Finally, we identify the characteristics that the algorithms associate with better performance.

We identify top performers as funds in the top quartile, given their types and vintage years. We find that machine learning algorithms can correctly classify a high proportion of funds compared to random investing and return-chasing, especially for all funds and buyout funds. However, the more complex models outperform the simpler logit models in their classifications.

We then create hypothetical portfolios following the algorithms' and the benchmark strategies' classifications. We find that the algorithms can generate economically large IRRs that exceed the third quartile values for the sample. Some algorithms also produce average IRRs almost ten percent higher than those observed by return chasing. Machine learning also creates a wider performance gap between predicted top-quartile and non-top-quartile funds than traditional benchmarks. While the algorithms do not show clearly higher average IRRs for venture funds compared to return chasing, they still outperform random investing. Most algorithms continue to perform well for buyout funds, during which period persistence declines.

We further identify a set of characteristics that the algorithms consider important in training the data. We find that most characteristics weighed more heavily by the algorithms are interaction terms that involve both qualitative and quantitative information about the GP's past, general private equity condition, and the specific fund focus. This speaks to the multi-dimensional part of private equity returns, as past performance alone cannot accurately predict the future.

As private equity becomes a mainstay in institutional investors' portfolios, one question is how we can identify top-performing funds ex-ante. To that end, we show that using machine learning to identify top-performing funds can have practical implications. Our results are based on limited structured data on fund and firm characteristics. We expect LPs, who have more information about the funds and their GPs, would yield more accurate predictions. Our results highlight the importance of information in predicting fund performance, particularly complex information that captures more than just past returns. Furthermore, even with machine learning, we cannot accurately predict private equity fund performance with limited hard information, at least to the extent of accuracy in other fields. This is consistent with the notion that analyzing complex information is valuable in assessing a GP's skill and the fund's performance.

REFERENCES

Ang, Y. Q., Chia, A., and Saghafian, S., 2021, "Using Machine Learning to Demystify Startups' Funding, Post-Money Valuation, and Success", *Innovative Technology at the Interface of Finance and Operations: Volume I*, 271-296.

Barber, B.M., and Yasuda, A., 2017, "Interim Fund Performance and Fundraising in Private Equity", *Journal of Financial Economics,* 124, 172-194.

Braun, R., Jenkinson, T., and Stoff, I., 2017, "How Persistent Is Private Equity Performance? Evidence from Deal-Level Data", *Journal of Financial Economics*, 123, 273-291.

Braun, R., Tamayo, B. F., López-de-Silanes, F., Phalippou, L., and Sigrist, N., 2023, "Limited Partners versus Unlimited Machines: Artificial Intelligence and the Performance of Private Equity Funds", Working Paper.

Brown, G. W., Ghysels, E., and Gredil, O. R., 2023, "Nowcasting Net Asset Values: The Case of Private Equity", *The Review of Financial Studies*, 36, 945-986.

Cavagnaro, D. R., Sensoy, B. A., Wang, Y., and Weisbach, M. S., 2019, "Measuring Institutional Investors' Skill at Making Private Equity Investments", *The Journal of Finance*, 74, 3089-3134.

Chung, J. W., 2012, "Performance Persistence in Private Equity Funds", Working Paper.

Chung, J. W., Sensoy, B. A., Stern, L., and Weisbach, M. S., 2012, "Pay for Performance from Future Fund Flows: The Case of Private Equity", *The Review of Financial Studies*, 25, 3259-3304.

Davenport, D., 2022, "Predictably Bad Investments: Evidence from Venture Capitalists", Working Paper.

Diller, C., and Kaserer, C., 2009, "What Drives Private Equity Returns? Fund Inflows, Skilled GPs, and/or Risk?", *European Financial Management*, 15, 643-675.

Erel, I., Stern, L. H., Tan, C., and Weisbach, M. S., 2021, "Selecting Directors Using Machine Learning", *The Review of Financial Studies*, 34, 3226-3264.

https://doi.org/10.55802/IJB.030(4).002

Ewens, M., and Rhodes-Kropf, M., 2015, "Is a VC Partnership Greater than the Sum of Its Partners?", *The Journal of Finance*, 70, 1081-1113.

Fragkiskos, A., Krasotkina, O., Spilker, H. D., III, and Wermers, R., 2025, "Private Equity Fund Performance: A Time-Series Approach", *Journal of Banking and Finance*, 177, 107470.

Gompers, P., and Lerner, J., 2000, "Money Chasing Deals? The Impact of Fund Inflows on Private Equity Valuation", *Journal of Financial Economics*, 55, 281-325.

Gu, S., Kelly, B., and Xiu, D., 2020, "Empirical Asset Pricing via Machine Learning", *The Review of Financial Studies*, 33, 2223-2273.

Gupta, A., and Van Nieuwerburgh, S., 2021, "Valuing Private Equity Investments Strip by Strip", *The Journal of Finance*, 76, 3255-3307.

Harris, R. S., Jenkinson, T., Kaplan, S. N., and Stucke, R., 2020, "Has Persistence Persisted in Private Equity? Evidence from Buyout and Venture Capital Funds", National Bureau of Economic Research.

Hochberg, Y. V., and Rauh, J. D., 2013, "Local Overweighting and Underperformance: Evidence from Limited Partner Private Equity Investments", *The Review of Financial Studies*, 26, 403-451.

Hochberg, Y. V., Ljungqvist, A., and Vissing-Jørgensen, A., 2014, "Informational Holdup and Performance Persistence in Venture Capital", *The Review of Financial Studies*, 27, 102-152.

Jenkinson, T., Sousa, M., and Stucke, R., 2013, “How Fair are the Valuations of Private Equity Funds?”, Working Paper.

Jones, C. M., and Rhodes-Kropf, M., 2003, "The Price of Diversifiable Risk in Venture Capital and Private Equity", Working Paper, Columbia University.

Kaniel, R., Lin, Z., Pelger, M., and Van Nieuwerburgh, S., 2022, "Machine-Learning the Skill of Mutual Fund Managers", National Bureau of Economic Research.

Kaplan, S. N., and Schoar, A., 2005, "Private Equity Performance: Returns, Persistence, and Capital Flows", *The Journal of Finance*, 60, 1791-1823.

Korteweg, A., and Sorensen, M., 2017, "Skill and Luck in Private Equity Performance", *Journal of Financial Economics*, 124, 535-562.

Lerner, J., Schoar, A., and Wongsunwai, W., 2007, "Smart Institutions, Foolish Choices: The Limited Partner Performance Puzzle", *The Journal of Finance*, 62, 731-764.

Li, K., Mai, F., Shen, R., and Yan, X., 2021, "Measuring Corporate Culture Using Machine Learning", *The Review of Financial Studies*, 34, 3265-3315.

Lyonnet, V., and Stern, L. H., 2025, "Machine Learning About Venture Capital Choices", Working Paper.

Phalippou, L., and Zollo, M., 2005, "What Drives Private Equity Fund Performance?", Financial Institutions Center, Wharton School, University of Pennsylvania.

Sensoy, B. A., Wang, Y., and Weisbach, M. S., 2014, "Limited Partner Performance and the Maturing of the Private Equity Industry", *Journal of Financial Economics*, 112, 320-343.

APPENDIX I

This appendix provides a description of the variables we use to train and test machine learning algorithms. Each algorithm, including Logit, uses the same set of variables to train the data, then tests the top-quartile classifications based on demeaned IRR in the validation dataset. This process is repeated 100 times. Table AI.1 is the list of variables used by each algorithm in the training data to classify fund performance. These variables are divided into fund, GP, and private equity levels. They capture different aspects and measures of funds, GPs, and general market conditions that can be associated with returns. All demeaned IRRs are computed by subtracting the average IRR of all funds of the same type and vintage year. Table AI.2 shows fund performance classifications for cutoffs that exceed 0.5. This table completes the cutoffs displayed in Table 2.
All variables are defined in Table 2 as well.

https://doi.org/10.55802/IJB.030(4).002

Table AI. 1

Variable Description

Variable	Definition
Fund-level:	
Lag IRR	IRR of the preceding fund of the same type raised by the same GP. If no preceding fund IRR is observed, the fund is deleted from the sample. These include all first-time funds.
Lag demeaned IRR	Demeaned IRR of the preceding fund of the same type raised by the same GP.
Fund size (in $ millions)	The size of the fund raised is measured in $ millions.
Fund growth (in $ millions)	The increase in fund size from the preceding fund of the same type.
Fund growth (%)	The percentage increase in fund size from the preceding fund of the same type.
IT indicator	An indicator variable that equals one if an IT-related industry appears in the fund's description, and zero otherwise.
Healthcare indicator	An indicator variable that equals one if the healthcare-related industry appears in the fund's description, and zero otherwise.
Fund region focus	The region focus of the fund is divided to US, Europe, and the rest of the world (ROTW).
GP-level:	
GP experience	The number of funds of the same fund type that the GP raised before the current fund.
GP age (in years)	The number of years it has been since the GP raised its first fund.
Fundraising gap (in years)	The number of years between the current and preceding fund of the same type.
Fund 1 vintage	The vintage year of the first fund raised by the GP.
GP average IRR	The average IRR of all funds of the same type raised by the same GP is weighted by fund size.
GP average demeaned IRR	The averaged demeaned IRR of all funds of the same type raised by the same GP weighed by fund size.
GP average 5-year IRR	The average IRR of all funds of the same type raised by the same GP within the last five years weighed by fund size.
GP average 5-year demeaned IRR	The averaged demeaned IRR of all funds of the same type raised by the same GP within the last five years weighed by fund size.
Same region indicator	An indicator variable that equals one if the fund region focus and GP location are the same.
GP location	GP location is divided to US, Europe, and the rest of the world (ROTW).
PE-level:	
HHI	Herfindahl-Hirschman Index of current vintage year is estimated using the size of funds in the current year.
Lag HHI	Herfindahl-Hirschman Index of the last vintage year was estimated using the size of funds raised in the prior year.
PE inflow (in $ millions)	The total inflow (measured in $ millions) to funds within the same fund type and vintage year.
Lag PE inflow (in $ millions)	The total size of funds (measured in $ millions) of the same type in the prior vintage year.

Table AI. 2
Fund Performance Classification

This table reports top-quartile (i.e., C_1) classification results for each algorithm with cutoffs exceeding 0.5. The rest of the cutoffs are reported in Table 2. All algorithms, column headings, and variable definitions are the same as those in Table 2.

Cutoff	Algorithm	All funds				Buyout funds				Venture funds			
		Predicted C_1	Accuracy	Sensitivity	Specificity	Predicted C_1	Accuracy	Sensitivity	Specificity	Predicted C_1	Accuracy	Sensitivity	Specificity
0.6	Logit 0	0.003	0.739	0.002	0.996	0.007	0.736	0.011	0.995	0.002	0.740	0.000	0.997
0.6	Logit	0.361	0.578	0.385	0.647	0.376	0.578	0.410	0.635	0.371	0.582	0.412	0.643
0.6	Glmnet	0.002	0.742	0.006	0.999	0.005	0.739	0.013	0.998	0.000	0.743	0.000	1.000
0.6	XGBoost	0.002	0.740	0.002	0.998	0.002	0.737	0.004	0.998	0.003	0.741	0.003	0.996
0.6	Random forest	0.014	0.738	0.019	0.988	0.021	0.735	0.039	0.984	0.022	0.734	0.028	0.979
0.6	GAM	0.003	0.739	0.001	0.996	0.007	0.735	0.009	0.994	0.007	0.737	0.002	0.991
0.7	Logit 0	0.003	0.739	0.000	0.996	0.004	0.739	0.010	0.999	0.001	0.742	0.000	0.999
0.7	Logit	0.361	0.578	0.385	0.647	0.376	0.578	0.410	0.635	0.371	0.582	0.412	0.643
0.7	Glmnet	0.001	0.742	0.005	1.000	0.002	0.739	0.008	1.000	0.000	0.743	0.000	1.000
0.7	XGBoost	0.000	0.741	0.000	1.000	0.000	0.737	0.000	1.000	0.001	0.742	0.000	0.999
0.7	Random forest	0.004	0.743	0.011	0.998	0.005	0.739	0.015	0.998	0.008	0.746	0.022	0.997
0.7	GAM	0.000	0.741	0.000	1.000	0.004	0.737	0.008	0.997	0.003	0.740	0.000	0.996
0.8	Logit 0	0.003	0.739	0.000	0.996	0.003	0.740	0.010	1.000	0.000	0.743	0.000	1.000
0.8	Logit	0.361	0.578	0.385	0.647	0.376	0.578	0.410	0.635	0.371	0.582	0.412	0.643
0.8	Glmnet	0.001	0.742	0.003	1.000	0.001	0.738	0.005	1.000	0.000	0.743	0.000	1.000
0.8	XGBoost	0.000	0.741	0.000	1.000	0.000	0.737	0.000	1.000	0.000	0.743	0.000	1.000
0.8	Random forest	0.000	0.741	0.000	1.000	0.000	0.737	0.000	1.000	0.001	0.743	0.002	1.000
0.8	GAM	0.000	0.741	0.000	1.000	0.002	0.739	0.008	1.000	0.001	0.742	0.000	0.999

0.9	Logit 0	0.002	0.740	0.000	0.998	0.002	0.739	0.008	1.000	0.000	0.743	0.000	1.000
0.9	Logit	0.361	0.578	0.385	0.647	0.376	0.578	0.410	0.635	0.371	0.582	0.412	0.643
0.9	Glmnet	0.000	0.741	0.001	1.000	0.000	0.737	0.000	1.000	0.000	0.743	0.000	1.000
0.9	XGBoost	0.000	0.741	0.000	1.000	0.000	0.737	0.000	1.000	0.000	0.743	0.000	1.000
0.9	Random forest	0.000	0.741	0.000	1.000	0.000	0.737	0.000	1.000	0.000	0.743	0.000	1.000
0.9	GAM	0.000	0.741	0.000	1.000	0.002	0.739	0.007	1.000	0.000	0.742	0.000	1.000

Appendix III

This appendix provides the main results for predicting bottom-quartile funds. Our aim is to provide a more comprehensive picture of whether using machine learning can help LPs make investment decisions, including identifying funds to avoid as well.

We follow the same process and benchmarks we used for top-quartile funds to classify funds in the bottom quartile of each fund type and vintage year. We call the funds predicted to be in the bottom quartile C_1’ and the rest C_0’. For return chasing, we identify funds to avoid as those with preceding funds in the bottom quartile, given their fund types and vintage years. To be brief, we report the main classification comparisons in Table A.II.1 (analogous to Table 2 in the paper).

Our results show that the algorithms can better identify bottom-quartile funds compared to the benchmarks. At a cutoff of 0.3, most algorithms show a higher proportion of predicted C_1’ funds and higher sensitivity for all funds, as well as venture and buyout funds separately.

Table A. II.1

Classification of Bottom-Quartile funds

Panel A: Machine-learning Classifications

Cutoff	Algorithm	All				Buyout				Venture Funds			
		Predicted C_1	Accuracy	Sensitivity	Specificity	Predicted C_1	Accuracy	Sensitivity	Specificity	Predicted C_1	Accuracy	Sensitivity	Specificity
0.1	Logit 0	0.983	0.272	0.989	0.019	0.967	0.292	0.990	0.042	0.986	0.277	0.985	0.014
0.1	Logit	0.360	0.585	0.390	0.650	0.398	0.565	0.427	0.613	0.379	0.551	0.371	0.619

		All				Buyout				Venture Funds			
Cutoff	Algorithm	Predicted C_1	Accuracy	Sensitivity	Specificity	Predicted C_1	Accuracy	Sensitivity	Specificity	Predicted C_1	Accuracy	Sensitivity	Specificity
0.1	Glmnet	0.991	0.267	0.994	0.011	0.975	0.284	0.992	0.031	0.997	0.270	0.992	0.002
0.1	XGBoost	0.990	0.267	0.993	0.011	0.991	0.270	0.993	0.009	0.984	0.280	0.988	0.017
0.1	Random forest	0.940	0.296	0.952	0.064	0.961	0.276	0.948	0.035	0.898	0.335	0.932	0.116
0.1	GAM	0.977	0.273	0.979	0.024	0.941	0.312	0.977	0.072	0.944	0.307	0.963	0.064
0.2	Logit 0	0.901	0.330	0.942	0.114	0.777	0.417	0.867	0.257	0.938	0.302	0.945	0.065
0.2	Logit	0.360	0.585	0.390	0.650	0.398	0.565	0.427	0.613	0.379	0.551	0.371	0.619
0.2	Glmnet	0.915	0.316	0.944	0.095	0.847	0.370	0.908	0.176	0.968	0.286	0.968	0.033
0.2	XGBoost	0.837	0.358	0.875	0.176	0.852	0.347	0.875	0.156	0.841	0.367	0.884	0.175
0.2	Random forest	0.693	0.438	0.749	0.328	0.697	0.415	0.710	0.309	0.661	0.483	0.770	0.380
0.2	GAM	0.795	0.384	0.843	0.223	0.731	0.439	0.821	0.303	0.748	0.434	0.839	0.287
0.3	Logit 0	0.196	0.703	0.309	0.844	0.299	0.655	0.418	0.745	0.114	0.712	0.190	0.912
0.3	Logit	0.360	0.585	0.390	0.650	0.398	0.565	0.427	0.613	0.379	0.551	0.371	0.619
0.3	Glmnet	0.156	0.699	0.224	0.868	0.221	0.689	0.333	0.821	0.050	0.708	0.057	0.953
0.3	XGBoost	0.514	0.543	0.611	0.520	0.520	0.529	0.599	0.509	0.551	0.520	0.633	0.479
0.3	Random forest	0.413	0.588	0.505	0.619	0.399	0.583	0.464	0.627	0.410	0.594	0.511	0.627
0.3	GAM	0.283	0.665	0.403	0.759	0.322	0.651	0.451	0.725	0.312	0.622	0.386	0.715
0.4	Logit 0	0.020	0.740	0.041	0.987	0.061	0.718	0.084	0.947	0.006	0.730	0.012	0.996
0.4	Logit	0.360	0.585	0.390	0.650	0.398	0.565	0.427	0.613	0.379	0.551	0.371	0.619
0.4	Glmnet	0.017	0.739	0.033	0.988	0.044	0.736	0.086	0.971	0.003	0.728	0.002	0.997
0.4	XGBoost	0.207	0.663	0.256	0.811	0.192	0.656	0.218	0.818	0.250	0.634	0.293	0.766
0.4	Random forest	0.183	0.683	0.243	0.839	0.184	0.665	0.213	0.828	0.211	0.663	0.268	0.811
0.4	GAM	0.063	0.730	0.105	0.952	0.106	0.720	0.173	0.918	0.074	0.703	0.088	0.932

This table reports the average classification of bottom-quartile funds' performance using both machine learning algorithms and benchmark investment strategies. C_1's funds are those with IRRs in the bottom quartile of the funds' vintage years and fund types. All algorithms, column headings, and variable definitions are the same as those in Table 2. Panel A shows results for the machine learning algorithms. Panel B reports similar results using random investing and return chasing. For return chasing, a fund is classified as C_1' if the preceding fund's IRR is in the bottom quartile for its fund type and vintage year.

Panel B: Benchmark Classifications

		All Funds				Buyout Funds				Venture Funds			
	Benchmark	Predicted C_1'	Accuracy	Sensitivity	Specificity	Predicted C_1'	Accuracy	Sensitivity	Specificity	Predicted C_1'	Accuracy	Sensitivity	Specificity
25%	Random Investing	0.25	0.63	0.25	0.75	0.25	0.63	0.25	0.75	0.25	0.63	0.25	0.75
25%	Return Chasing	0.25	0.64	0.32	0.84	0.25	0.63	0.28	0.83	0.25	0.65	0.36	0.85

Table 1.
Summary Statistics of fund and GP characteristics

This table shows summary statistics of fund and GP characteristics in our sample. Column N shows the number of observations for each variable. Columns Q1 and Q3 each represents first and third quartile values, respectively. Fund region focus, GP location, and demeaned IRRs are also used by the algorithms but are not included in the summary statistics table. Table AI.1 includes a detailed description of each variable used by the algorithms.

	All Funds					Buyout Funds					Venture Funds				
	N	Mean	Median	Q1	Q3	N	Mean	Median	Q1	Q3	N	Mean	Median	Q1	Q3
IRR	1,402	13.94	10.60	3.20	19.60	776	14.73	12.80	7.65	20.50	626	12.96	6.45	-2.10	17.50
Lag IRR	1,402	18.57	12.80	4.90	24.60	776	18.92	15.95	9.15	25.50	626	18.15	8.00	-10	21.50
Fund size (in $ millions)	1,402	951	341	150	806	776	1,498	629	290	1,552	626	274	180	87	354
Fund growth (in $ millions)	1,402	345.51	102	7	325	776	578.81	208.06	52.63	691.50	626	56.31	44.92	-3.84	122
Fund growth (%)	1,402	1.22	0.61	0.05	1.36	776	1.38	0.73	0.20	1.50	626	1.02	0.44	-0.04	1.16
IT indicator	1,402	0.35	0	0	1	776	0.16	0	0	0	626	0.59	1	0	1
Healthcare indicator	1,402	0.42	0	0	1	776	0.35	0	0	1	626	0.52	1	0	1
Fundraising gap (in years)	1,402	4	3	2	5	776	4	4	2	5	626	3	3	2	4
GP experience	1,402	3.32	2	1	4	776	3.06	2	1	4	626	3.64	3	1	4
GP age (in years)	1,402	9.38	8	5	13	776	9.26	8	5	13	626	9.53	8	4	14
Fund 1 vintage	1,402	1994	1994	1988	1999	776	1995	1995	1990	1999	626	1993	1993	1986	1998
GP avg IRR	1,402	19.49	15.50	8.60	24.60	776	19.95	17.53	12.29	25.15	626	18.93	11.05	3.87	23.25
GP 5-year IRR	1,402	19.65	15.78	8.40	25.00	776	20.11	17.88	12.29	25.84	626	19.08	19.08	11.40	23.8
Same region indicator	1,402	0.93	1	1	1	776	0.92	1	1	1	626	0.95	1	1	1
HHI	1,402	0.03	0.02	0.01	0.03	776	0.03	0.03	0.02	0.04	626	0.02	0.01	0.01	0.02
Prior year HHI	1,402	0.03	0.02	0.01	0.03	776	0.04	0.03	0.02	0.04	626	0.02	0.01	0.01	0.02
PE inflow (in $ millions)	1,402	87,608	51,910	35,322	128,561	776	128,554	105,994	60,896	216,929	626	36,850	44,019	18,014	49,639
Prior year PE inflow (in $ millions)	1,402	75,983	51,239	23,742	71,376	776	111,597	66,917	51,910	149,357	626	31,835	32,757	14,805	49,078

Table 2.

Classification of Top Quartile Funds

This table reports the average classification of funds' performance using both machine learning algorithms and benchmark investment strategies. The algorithms we use are Logit, GLMnet, XGBoost, random forest, and generalized additive model (GAM). We use two Logit models. The first (Logit 0) only includes the preceding fund demeaned IRR. IRRs are demeaned by subtracting the average IRR of the same fund type and vintage year. The second (Logit) uses the full range of variables, as the other algorithms do. For each algorithm, we randomly select 60% of the full sample to train the algorithms and use the remaining 40% for validation. We repeat the process 100 times, then take the average classification across the 100 cycles. Panel A reports results for each algorithm for all funds, buyout funds, and venture funds. Column Cutoff shows the probability for each fund to be in class 1 (C_1) given its fund type and vintage year. We define C_1 funds as those with top-quartile demeaned IRRs given their vintage years and fund types. The rest of the funds are classified as C_0. If cutoff=0.5, then a fund is classified as C_1 if the average probability of the fund being in the top quartile exceeds 50%. Column Predicted C_1 shows the percentage of true top-quartile funds identified by each algorithm in the validation set. Column Accuracy shows the proportions of correctly classified funds (both fund classes) given their actual performance. Sensitivity and Specificity show the proportion of funds correctly classified as C_1 and C_0 based on the actual performance, respectively. Panel B reports similar results using benchmark strategies: random investing and return chasing. 25% represents the performance percentile cutoffs for classifications. For return chasing, this 25% means that if the preceding fund's IRR is in the top quartile for its type and vintage year, we classify the follow-on fund as a C_1 fund.

Panel A: machine-learning classifications

		All				Buyout				Venture Funds			
Cutoff	Algorithm	Predicted C_1	Accuracy	Sensitivity	Specificity	Predicted C_1	Accuracy	Sensitivity	Specificity	Predicted C_1	Accuracy	Sensitivity	Specificity
0.1	Logit 0	1.000	0.259	1.000	0.000	0.996	0.266	0.998	0.004	1.000	0.257	1.000	0.000
0.1	Logit	0.361	0.578	0.385	0.647	0.376	0.578	0.410	0.635	0.371	0.582	0.412	0.643
0.1	Glmnet	0.997	0.260	0.996	0.003	0.981	0.269	0.976	0.017	1.000	0.258	1.000	0.000

https://doi.org/10.55802/IJB.030(4).002

		All				Buyout				Venture Funds			
Cutoff	Algorithm	Predicted C_1	Accuracy	Sensitivity	Specificity	Predicted C_1	Accuracy	Sensitivity	Specificity	Predicted C_1	Accuracy	Sensitivity	Specificity
0.1	XGBoost	0.994	0.262	0.995	0.007	0.994	0.268	0.997	0.007	0.988	0.263	0.989	0.012
0.1	Random forest	0.953	0.287	0.963	0.051	0.945	0.305	0.975	0.066	0.944	0.293	0.961	0.062
0.1	GAM	0.987	0.267	0.990	0.014	0.983	0.271	0.984	0.018	0.959	0.284	0.973	0.046
0.2	Logit 0	0.993	0.264	0.996	0.009	0.845	0.360	0.893	0.172	1.000	0.257	0.999	0.000
0.2	Logit	0.361	0.578	0.385	0.647	0.376	0.578	0.410	0.635	0.371	0.582	0.412	0.643
0.2	Glmnet	0.909	0.324	0.949	0.105	0.820	0.399	0.917	0.216	0.988	0.261	0.983	0.011
0.2	XGBoost	0.900	0.319	0.923	0.108	0.890	0.338	0.935	0.125	0.855	0.338	0.879	0.154
0.2	Random forest	0.698	0.448	0.781	0.332	0.686	0.464	0.791	0.350	0.695	0.442	0.769	0.331
0.2	GAM	0.814	0.381	0.876	0.209	0.724	0.430	0.797	0.301	0.840	0.342	0.858	0.166
0.3	Logit 0	0.059	0.725	0.085	0.950	0.224	0.679	0.319	0.809	0.029	0.730	0.033	0.973
0.3	Logit	0.361	0.578	0.385	0.647	0.376	0.578	0.410	0.635	0.371	0.582	0.412	0.643
0.3	Glmnet	0.166	0.693	0.232	0.857	0.263	0.675	0.393	0.780	0.020	0.731	0.017	0.979
0.3	XGBoost	0.525	0.522	0.594	0.500	0.533	0.516	0.600	0.489	0.534	0.512	0.596	0.488
0.3	Random forest	0.403	0.590	0.488	0.627	0.394	0.589	0.474	0.632	0.398	0.598	0.498	0.636
0.3	GAM	0.273	0.665	0.381	0.765	0.304	0.654	0.426	0.738	0.247	0.647	0.302	0.771
0.4	Logit 0	0.011	0.738	0.014	0.990	0.051	0.739	0.102	0.967	0.006	0.738	0.002	0.993
0.4	Logit	0.361	0.578	0.385	0.647	0.376	0.578	0.410	0.635	0.371	0.582	0.412	0.643
0.4	Glmnet	0.021	0.735	0.029	0.981	0.050	0.727	0.077	0.959	0.002	0.741	0.000	0.998
0.4	XGBoost	0.205	0.655	0.232	0.804	0.196	0.663	0.238	0.818	0.209	0.649	0.228	0.797
0.4	Random forest	0.174	0.686	0.230	0.845	0.192	0.673	0.248	0.827	0.191	0.664	0.225	0.820
0.4	GAM	0.051	0.721	0.059	0.952	0.078	0.725	0.125	0.939	0.043	0.725	0.050	0.960
0.5	Logit 0	0.005	0.738	0.005	0.995	0.017	0.734	0.028	0.987	0.004	0.739	0.001	0.994

		All				Buyout				Venture Funds			
Cutoff	Algorithm	Predicted C_1	Accuracy	Sensitivity	Specificity	Predicted C_1	Accuracy	Sensitivity	Specificity	Predicted C_1	Accuracy	Sensitivity	Specificity
0.5	Logit	0.361	0.578	0.385	0.647	0.376	0.578	0.410	0.635	0.371	0.582	0.412	0.643
0.5	Glmnet	0.006	0.740	0.008	0.995	0.015	0.733	0.021	0.987	0.001	0.742	0.000	0.999
0.5	XGBoost	0.012	0.737	0.014	0.989	0.022	0.728	0.026	0.979	0.019	0.733	0.020	0.981
0.5	Random forest	0.057	0.722	0.074	0.949	0.066	0.721	0.097	0.944	0.070	0.711	0.077	0.932
0.5	GAM	0.012	0.734	0.011	0.987	0.028	0.730	0.040	0.976	0.015	0.735	0.015	0.985

Panel B: benchmark classifications

		All				Buyout				Venture Funds			
Benchmark		Predicted C_1	Accuracy	Sensitivity	Specificity	Predicted C_1	Accuracy	Sensitivity	Specificity	Predicted C_1	Accuracy	Sensitivity	Specificity
25%	Random investing	0.25	0.63	0.25	0.75	0.25	0.63	0.25	0.75	0.25	0.63	0.25	0.75
25%	Return chasing	0.25	0.69	0.37	0.78	0.25	0.69	0.40	0.78	0.25	0.68	0.32	0.78

Figure 1.
Comparisons of Assigned Probabilities Across Algorithms

This figure compares the assigned probabilities of true C_1 (i.e., top-quartile) funds and true C_0 funds from all 100 random test datasets. Each algorithm is coded in a different color, with the order of Logit 0, Logit, GLMnet, XGBoost, random forest (Forest), and GAM. Each algorithm has two boxes. The *High* box represents the probability of assigning a true top-quartile fund as a C_1 fund. The *Low* box represents the probability of assigning a true non-top fund as a C_0 fund. If the *High* box is shifted higher than the *Low* box, then the algorithm assigns higher probabilities of being C_1 funds to true top-quartile funds. *All* shows comparisons for all funds, while *Buyout* and *VC* separate all funds to buyout and venture funds.

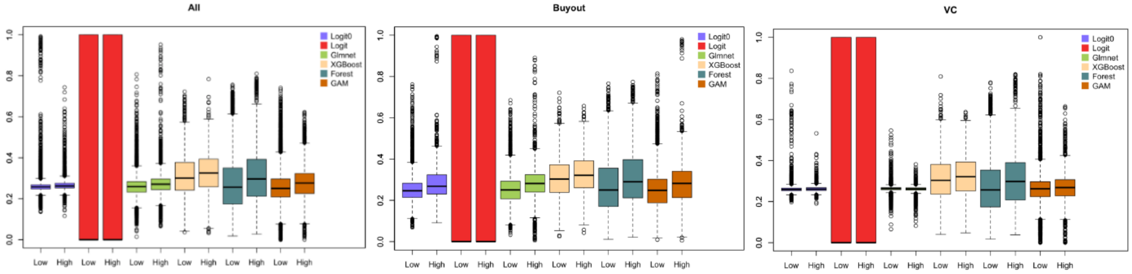

Figure 2.
Distribution of P-Values Across Algorithms

The figure shows box plots of the distribution of P-values from 100 cycles of the test data. Each color represents a different algorithm. They are listed in the order of Logit0, Logit, Glmnet, XGBoost, random forest (Forest), and GAM. For each of the 100 random splits, we conduct a two-sample t-test of assigned probabilities between true C_1 and C_0 observations in the test data. The p-values for all funds (ALL), buyout funds (buyout), and venture capital funds (VC) are reported separately.

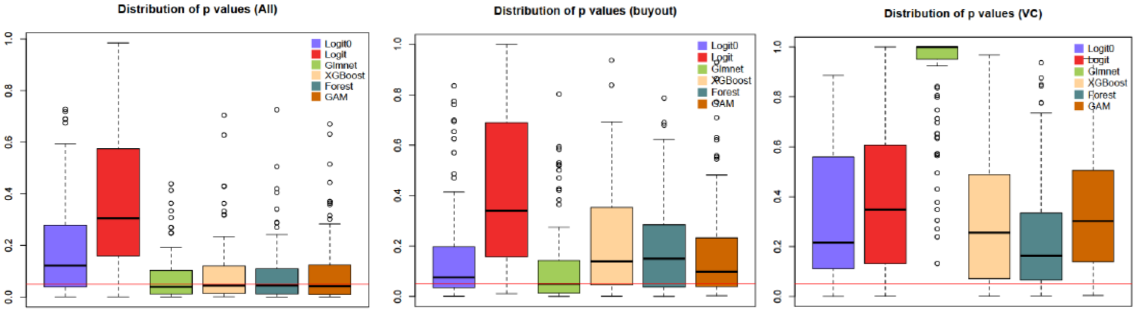

Table 3.

Hypothetical Average Returns Using Actual IRRs

This table reports average IRRs if hypothetical investments were made based on the classification of different algorithms and benchmarks. For each cycle, we calculate the average IRR of funds classified as C_1 (i.e., top quartile) and C_0. We then take the average of those averages across 100 cycles. Column Average IRR shows the average of observed IRRs across 100 cycles. For each cycle, we also compute the demeaned IRR by subtracting the mean IRR of all funds of the same type and vintage year from the actual IRR. Column Demeaned Average reports the average across 100 cycles. Columns C_1 and C_0 each represent C_1 and C_0 funds, respectively. Panel A shows results for all funds. Panels B and C show results for buyout funds and venture funds separately. The optimal cutoff for the algorithms is the cutoff that maximizes the average IRR of C_1. For the two benchmarks, the cutoff points remain 25%. For return chasing, C_1 funds are the follow-on funds of top-quartile funds, given their fund types and vintage years.

Panel A: All funds

Algorithm	Optimal Cutoff	Average IRR		Demeaned IRR	
		C_1	C_0	C_1	C_0
Logit 0	0.4	53	13.5	36.5	-0.1
Logit	0.1	14.8	13.5	1.1	-0.3
Glmnet	0.8	22	14.1	8	0.5
XGBoost	0.4	18.9	12.9	4	-0.5
Random Forest	0.8	36	13.8	19	-0.7
GAM	0.5	39.7	13.5	29.2	-0.2
Benchmark					
Random	0.25	13.94	13.94	0	0
Return-chasing	0.25	19.09	10.71	5.01	-1.75

Panel B: Buyout Funds

Algorithm	Optimal Cutoff	Average IRR		Demeaned IRR	
		C_1	C_0	C_1	C_0
Logit 0	0.4	25	14.5	10.5	-0.6
Logit	0.1	16.2	14.6	1.1	-0.4
Glmnet	0.7	27.9	14.6	13	-0.5
XGBoost	0.6	26.1	14.6	9.9	-0.2
Random Forest	0.7	33.4	14.7	16.1	-0.4
GAM	0.9	24.1	14.8	8.7	-0.5
Benchmark					
Random	0.25	14.73	14.73	0	0
Return-chasing	0.25	18.11	12.8	3.06	-1.63

Panel C: Venture Funds

Algorithm	Optimal Cutoff	Average IRR		Demeaned IRR	
		C_1	C_0	C_1	C_0
Logit 0	0.3	42.2	11.6	28	-0.6
Logit	0.1	13.3	12	0.8	0
Glmnet	0.2	12.6	5.7	0.4	-3.3
XGBoost	0.8	40.8	10.3	4.4	-0.6
Random Forest	0.3	17.3	9.5	4.6	-2.5
GAM	0.4	30	11.7	19.4	-0.6
Benchmark					
Random	0.25	12.96	12.96	0	0
Return-chasing	0.25	20.45	8.05	7.71	-1.91

Table 4.

Hypothetical Average Returns Using Average Fund Class IRRs

This table reports average IRRs if hypothetical investments were made based on the classification of different algorithms and benchmark strategies. The IRR for each fund is the average IRR of the fund's class instead of the actual IRR as reported in Table 3. Specifically, if the fund's true performance is in C_1, then we assign the average of all top-quartile funds of the same type and vintage year as the fund's IRR. We do the same for C_0 funds. All columns are defined in Table 3.

Panel A: All funds

Algorithm	Optimal Cutoff	Average IRR		Demeaned IRR	
		C_1	C_0	C_1	C_0
Logit 0	0.5	20.2	13.9	0.2	0.1
Logit	0.1	14.9	13.5	0.8	-0.1
Glmnet	0.9	39.9	13.8	21.4	0.5
XGBoost	0.6	18.1	13.9	2.0	0.4
Random Forest	0.8	39.9	13.7	21.4	0.2
GAM	0.3	17.6	12.6	3.1	-0.9
Benchmark					
Random	25%	13.94	13.94	0	0
Return-chasing	25%	16.97	11.48	2.88	-0.98

Panel B: Buyout Funds

Algorithm	Optimal Cutoff	Average IRR		Demeaned IRR	
		C_1	C_0	C_1	C_0
Logit 0	0.8	33.3	14.7	16.3	-0.5
Logit	0.1	15.6	14.3	0.5	-0.6
Glmnet	0.7	33.3	14.7	16.3	-0.4
XGBoost	0.7	33.3	14.9	-5.8	-0.1
Random Forest	0.7	25.0	14.5	6.8	-0.3
GAM	0.9	33.3	14.8	16.3	-0.5
Benchmark					
Random	25%	14.73	14.73	0	0
Return-chasing	25%	17.08	13.19	2.02	-1.24

Panel C: Venture Funds

Algorithm	Optimal Cutoff	Average IRR		Demeaned IRR	
		C_1	C_0	C_1	C_0
Logit 0	0.4	20.8	12.5	-7.4	0.6
Logit	0.1	13.9	12.4	1.8	0.3
Glmnet	0.2	12.9	7.7	0.7	-1.9
XGBoost	0.8	47.8	13.1	-8.2	2.3
Random Forest	0.4	16.4	12.1	2.0	0.4
GAM	0.5	20.0	12.5	-0.7	0.6
Benchmark					
Random	25%	12.96	12.96	0	0
Return-chasing	25%	16.81	9.32	4.07	-0.64

Table 5.
Hypothetical Returns during post-2000 Sample Period

This table reports average IRRs of C_1 and C_0 funds using the post-2000 (including 2000) sample period. We rerun the same tests reported in Tables 3 and 4 during this sample period. To be brief, we only report the demeaned average IRRs. Results using non-demeaned IRRs are not meaningfully different. Column Average Actual IRR reports average IRRs of predicted C_1 and C_0 funds using their actual demeaned IRRs. Column Average Class IRR reports average IRRs of predicted C_1 and C_0 funds using the average demeaned IRR of the entire C_1 and C_0 funds. All other variables are defined in the previous tables. Panel A shows results for all funds. Panels B and C show results for buyout funds and venture funds separately.

Panel A: All funds

Algorithm	Average Actual Demeaned IRR			Average Class Demeaned IRR		
	Optimal Cutoff	C_1	C_0	Optimal Cutoff	C_1	C_0
Logit 0	0.3	4.9	-0.4	0.3	2	-0.2
Logit	0.1	0.8	-0.6	0.1	0.6	-0.3
Glmnet	0.4	5.6	-0.2	0.6	3.7	-0.1
XGBoost	0.6	2.2	-0.2	0.7	2.9	-0.1
Random Forest	0.7	11.6	-0.2	0.7	11.3	-0.1
GAM	0.3	2	-1	0.3	1.6	-0.7
Benchmark						
Random	25%	0	0	25%	0	0
Return-chasing	25%	1.73	-0.22	25%	0.71	-0.09

Panel B: Buyout Funds

Algorithm	Average Actual Demeaned IRR			Average Class Demeaned IRR		
	Optimal Cutoff	C_1	C_0	Optimal Cutoff	C_1	C_0
Logit 0	0.5	12.3	0.1	0.9	15.2	0.1
Logit	0.5	1.1	-0.2	0.1	0.5	-0.2
Glmnet	0.4	6.6	0	0.8	5.7	0.1
XGBoost	0.9	19	0.3	0.9	15.2	0.1
Random Forest	0.8	44.9	0.3	0.8	15.2	0.1
GAM	0.9	8.5	0.3	0.9	14.8	0.1
Benchmark						
Random	25%	0	0	25%	0	0
Return-chasing	25%	2.07	-0.31	25%	1.03	-0.15

Panel C: Venture Funds

Algorithm	Average Actual Demeaned IRR			Average Class Demeaned IRR		
	Optimal Cutoff	C_1	C_0	Optimal Cutoff	C_1	C_0
Logit 0	0.2	-0.5	-0.9	0.2	-0.5	1.5
Logit	0.3	-0.1	-0.8	0.6	-0.1	-0.6
Glmnet	0.1	-0.5	-1.3	0.1	-0.5	-0.5
XGBoost	0.6	2.1	-0.5	0.3	-0.1	-0.6
Random Forest	0.4	1	-0.9	0.4	0.9	-0.8
GAM	0.5	1.9	-0.6	0.5	-1	-0.4
Benchmark						
Random	25%	0	0	25%	0	0
Return-chasing	25%	1.12	-0.12	25%	0.14	-0.01

Table 6
Classifications Accuracy Without Potential Look-Ahead Bias

This table reports accuracy, sensitivity, and specificity without potential look-ahead bias. We train the algorithms using data from 1990 up to time n, then validate the algorithms using data from the following year only. We started our validation year in 2007 to ensure a large enough training set. We also remove variables that may not be available at the time of fundraising. These include variables related to lag IRR, fund size, and fund growth. Logit 0 is removed from this table as a result of removing lag IRR. We only use a cutoff of 0.2 to be brief.

https://doi.org/10.55802/IJB.030(4).002

Algorithm	Year	All funds			Buyout funds			Venture funds		
		Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity
Logit	2007	29.3	91.2	8.1	45.8	42.9	46.8	34	100	10.8
	2008	30.7	96.6	8.2	47	64.7	40.8	27.1	100	2.8
	2009	28.3	92.9	5.1	28.6	77.8	11.5	27.8	80	7.7
	2010	28.6	88.2	6.5	29.3	81.8	10	31.8	100	6.2
	2011	27.5	100	1.7	26.5	100	0	35.5	100	13
Glmnet	2007	25.6	100	0	25.3	100	0	26	100	0
	2008	25.4	100	0	25.8	100	0	25	100	0
	2009	26.4	100	0	25.7	100	0	27.8	100	0
	2010	27	100	0	26.8	100	0	27.3	100	0
	2011	26.2	100	0	26.5	100	0	25.8	100	0
XGBoost	2007	53.4	41.2	57.6	57.8	47.6	61.3	52	15.4	64.9
	2008	54.4	48.3	56.5	60.6	41.2	67.3	64.6	58.3	66.7
	2009	54.7	57.1	53.8	68.6	66.7	69.2	50	40	53.8
	2010	55.6	35.3	63	43.9	27.3	50	72.7	33.3	87.5
	2011	40	57.1	33.9	53.1	76.9	44.4	48.4	50	47.8
Random Forest	2007	60.2	23.5	72.7	62.7	33.3	72.6	56	30.8	64.9
	2008	68.4	13.8	87.1	68.2	23.5	83.7	64.6	41.7	72.2
	2009	69.8	35.7	82.1	74.3	44.4	84.6	61.1	60	61.5
	2010	71.4	17.6	91.3	73.2	27.3	90	63.6	0	87.5
	2011	68.8	28.6	83.1	61.2	23.1	75	61.3	37.5	69.6
GAM	2007	31.6	82.4	14.1	33.7	57.1	25.8	30	92.3	8.1
	2008	36	96.6	15.3	33.3	88.2	14.3	33.3	100	11.1
	2009	28.3	92.9	5.1	31.4	100	7.7	16.7	60	0
	2010	28.6	88.2	6.5	34.1	100	10	27.3	83.3	6.2
	2011	30	100	5.1	26.5	100	0	38.7	100	17.4

Table 7

Variance of Top-Quartile Fund Returns

This table reports the variances of returns of top-quartile funds identified by machine learning algorithms and return chasing. The cutoff for each algorithm is 0.2 for easier comparison with return chasing. Column IRR reports the variance of C_1 fund's IRR, and Column Demeaned IRR reports the variance of C_1 funds' IRRs demeaned by the average IRR of funds of the same type and vintage year. As a benchmark, return-chasing reports the variance of returns of funds whose preceding fund is in the top quartile within its type and vintage year.

Algorithm	IRR			Demeaned IRR		
	All	Buyout	VC	All	Buyout	VC
Logit 0	574.8	254.1	1032.7	450.4	225.6	789.1
Logit	571	250.4	867.9	432.8	212.6	627.4
Glmnet	615.5	286	1047.5	481.8	252.6	800.4
XGBoost	603.8	264.8	1105.09	473.8	230.1	837.9
Random Forest	679	256.1	1238.3	526.5	215.8	936.1
GAM	256.1	267.8	1163.7	490	237.6	868.9
Benchmark						
Return-chasing	1346.3	228.3	2816.9	1085.8	184.4	2270.2

https://doi.org/10.55802/IJB.030(4).002

Table 8

Characteristics that Affect Fund Performance

This table reports the top 20 characteristics used in classifying fund performance into the top quartile by Glmnet and random forest. The variables are ranked based on their importance in each algorithm, with 1 being the most important. Importance for Glmnet is determined by the frequency of the variable being selected (Shown in Column Selection frequency) across 100 cycles. The importance of random forest is determined by the mean importance score (shown in Column Importance Score) across cycles. "×" indicates interaction between variables. The definition of each variable is in Table AI.1.

Ranking	Glmnet	Selection frequency	Random forest	Importance Score
1	Buyout indicator × lag demeaned IRR	100	Vintage × Lag demeaned IRR	4.972
2	Same region indicator × healthcare indicator	98	Lag demeaned IRR	4.945
3	PE inflow × fund region focus-Europe	91	Lag demeaned IRR × same region indicator	4.48
4	GP age × GP average demeaned IRR	86	GP age × GP average demeaned IRR	4.289
5	Fund size × fund region focus-ROTW	84	Vintage × GP average 5-year demeaned IRR	4.066
6	Fund growth (%) × HHI	79	GP average 5-year demeaned IRR	4.021
7	Lag PE inflow × GP location-Europe	73	Fund size × PE inflow	3.84
8	Fund growth (%) × fund region focus-Europe	69	Lag demeaned IRR × fundraising gap (in years)	3.727
9	GP age × GP location-Europe	62	GP average demeaned IRR^2	3.683
10	HHI × fund region focus-ROTW	62	Fund size × HHI	3.683
11	PE inflow × GP location-ROTW	61	GP average 5-year demeaned IRR × same region indicator	3.679
12	Buyout indicator × GP average 5-year demeaned IRR	59	Lag demeaned IRR × GP average 5-year demeaned IRR	3.613
13	Lag demeaned IRR × Healthcare indicator	59	Fund growth (in $ millions)^2	3.543

Ranking	Glmnet	Selection frequency	Random forest	Importance Score
14	Fundraising gap (in years) × GP location-ROTW	57	Lag demeaned IRR × GP age	3.532
15	GP age × healthcare indicator	54	GP average 5-year demeaned IRR × fundraising gap (in years)	3.5
16	healthcare indicator × GP location-ROTW	52	Lag PE inflow × Lag demeaned IRR	3.437
17	Buyout indicator × healthcare indicator	48	Fund size × lag PE inflow	3.411
18	Lag demeaned IRR	41	GP average 5-year demeaned IRR × GP age	3.404
19	Fund growth (in $ millions) × Same region indicator	38	GP average 5-year demeaned IRR × fund growth (%)	3.393
20	Fund size × same region indicator	36	GP age × fund growth (%)	3.39